EMERGING TECH CONFERENCE – Edge Intelligence

Volume 02, 2023, Page 151 – 157

Proceedings of Emerging Tech Conference:
Edge Intelligence 2023

MORCIC: Model Order Reduction Techniques for Electromagnetic Models of Integrated Circuits

Dimitrios Garyfallou[1], Athanasios Stefanou[2], Christos Giamouzis[1], Moschos Antoniadis[2], Georgios Chararas[2], Konstantinos Chatzis[2], Dimitris Samaras[2], Rafaela Themeli[2], Anastasios Michailidis[2], Vasiliki Gogolou[2], Nikos Zachos[2], Nestor Evmorfopoulos[1,] Thomas Noulis[2], Vasilis F. Pavlidis[2], Alkiviadis Hatzopoulos[2], Elpida Chatzineofytou[3], and Yiannis Moisiadis[3]

[1] *Dept. of Electrical and Computer Engineering, University of Thessaly, Volos, Greece*
{digaryfa, cgiamouzis, nestevmo}@e-ce.uth.gr
[2] *Aristotle University of Thessaloniki, Thessaloniki, Greece*
{tnoul, vpavlid, alkis}@auth.gr
[3] *ANSYS-Hellas, Athens, Greece*
{elpida.chatzineofytou, yiannis.moisiadis}@ansys.com

Abstract

Model order reduction (MOR) is crucial for the design process of integrated circuits. Specifically, the vast amount of passive RLCk elements in electromagnetic models extracted from physical layouts exacerbates the extraction time, the storage requirements, and, most critically, the post-layout simulation time of the analyzed circuits. The MORCIC project aims to overcome this problem by proposing new MOR techniques that perform better than commercial tools. Experimental evaluation on several analog and mixed-signal circuits with millions of elements indicates that the proposed methods lead to ×5.5 smaller ROMs while maintaining similar accuracy compared to golden ROMs provided by ANSYS RaptorX™.

1 Introduction

Electromagnetic model extraction plays a key role in the design and analysis of integrated circuits. The extracted models are simulated to accurately predict the behavior of the passive elements of the design. Model order reduction (MOR) can reduce the complexity of RLCk models with many elements (>1M) and ports (>10), while retaining an accurate approximation of the input and output behavior of the circuit [1, 2]. Therefore, the simulation time of complex systems can be radically decreased by constructing reduced-order models (ROMs) of smaller dimensions that preserve the essential characteristics of the original models.

MOR methods are distinguished into two main categories. Moment matching (MM) techniques [1] are preferred due to their computational efficiency. However, they rely on an ad hoc selection of the number of moments, which correlates the final ROM size with the number of ports. On the other hand, techniques based on balanced truncation (BT) [2] offer reliable bounds for the approximation error and have no fundamental limitation to the number of ports they can handle, resulting in more compact ROMs. Nevertheless, BT applies only to small-scale models since it involves the computationally expensive

solution of Lyapunov equations [2].

In this work, appropriate performance improvements are explored to overcome the main drawback of the conventional BT method. To this end, we adopt an efficient low-rank technique based on the extended Krylov subspace (EKS) for solving the Lyapunov equations. The proposed approach can be integrated into industrial extraction tools, such as the ANSYS Rap torX™ [3], to obtain more compact ROMs of large-scale multi-port RLCk models.

2 Background

Consider the modified nodal analysis (MNA) description [4] of an n-node, m-branch (inductive), p-input, and q-output RLCk circuit in the time domain:

$$\begin{pmatrix} G_n & E \\ -E^T & 0 \end{pmatrix}\begin{pmatrix} v(t) \\ i(t) \end{pmatrix} + \begin{pmatrix} G_n & 0 \\ 0 & M \end{pmatrix}\begin{pmatrix} \dot{v}(t) \\ i(t) \end{pmatrix} = \begin{pmatrix} B_i \\ 0 \end{pmatrix} u(t), \quad y(t) = \begin{pmatrix} L_1 & 0 \end{pmatrix}\begin{pmatrix} v(t) \\ i(t) \end{pmatrix} \tag{1}$$

where $G_n \in \mathbb{R}^{n\times n}$ (node conductance matrix), $C_n \in \mathbb{R}^{n\times n}$ node capacitance matrix), $M \in \mathbb{R}^{m\times n}$ (branch inductance matrix), $E \in \mathbb{R}^{n\times m}$ (node-to-branch incidence matrix), $V \in \mathbb{R}^n$ (vector of node voltages), $i \in \mathbb{R}^m$ (vector of inductive branch currents), $u \in \mathbb{R}^p$ (vector of input excitations), $B_1 \in \mathbb{R}^{n\times p}$ (input-to-node connectivity matrix), $y \in \mathbb{R}^q$ (vector of output measurements), and $L1_l \in \mathbb{R}^{q\times n}$ (node-to-output connectivity matrix). Moreover, we denote $\dot{v}(t) \equiv \frac{dv(t)}{dt}$and $i(t) \equiv \frac{di(t)}{dt}$. If we now define the model order as $N \equiv n + m$, the state vector as $x(t) \equiv \begin{pmatrix} v(t) \\ i(t) \end{pmatrix}$, and also:

$$G \equiv -\begin{pmatrix} G_n & E \\ -E^T & 0 \end{pmatrix}, \qquad C \equiv \begin{pmatrix} G_n & 0 \\ 0 & M \end{pmatrix}, \qquad B \equiv \begin{pmatrix} B_i \\ 0 \end{pmatrix}, \qquad L \equiv \begin{pmatrix} L_1 & 0 \end{pmatrix}$$

then Eq. (1) can be written in the generalized state-space form, or so-called descriptor form:

$$C\frac{dx(t)}{dt} = Gx(t) + Bu(t), \quad y(t) = Lx(t) \tag{2}$$

The objective of MOR is to produce an equivalent ROM:

$$\tilde{C}\frac{d\tilde{x}(t)}{dt} = \tilde{G}\tilde{x}(t) + \tilde{B}u(t), \quad \tilde{y}(t) = \tilde{L}\tilde{x}(t) \tag{3}$$

Where $\tilde{G}, \tilde{C} \in \mathbb{R}^{rxr}$, $\tilde{B} \in \mathbb{R}^{rxp}$, $\tilde{L} \in \mathbb{R}^{qxr}$, the reduced order r << N, and the output error is bounded as $\|\tilde{y}(t) - y(t)\|_2 < \varepsilon\|u(t)\|_2$ for given u(t) and small ε. The output error bound can be expressed in the frequency domain as $\|\tilde{y}(s) - y(s)\|_2 < \varepsilon\|u(s)\|_2$ via Plancherel's theorem [5]. If

$$H(s) = L(sC - G)^{-1}B, \qquad \tilde{H}(s) = \tilde{L}(s\tilde{C} - \tilde{G})^{-1}\tilde{B}$$

are the transfer functions of the original model and the ROM, the corresponding output error is:

$$\|\tilde{\mathrm{y}}(\mathrm{s}) - \mathrm{y}(\mathrm{s})\|_2 = \left\|\tilde{\mathrm{H}}(\mathrm{s})\mathrm{u}(\mathrm{s}) - \mathrm{H}(\mathrm{s})\mathrm{u}(\mathrm{s})\right\|_2 \leq \left\|\tilde{\mathrm{H}}(\mathrm{s}) - \mathrm{H}(\mathrm{s})\right\|_\infty \|\mathrm{u}(\mathrm{s})\|_2$$

Where $\|.\|_\infty$ is the $\mathcal{L}_2$ matrix norm or $\mathcal{H}_\infty$ norm of a rational transfer function. Thus, to bound this error, we need to bound the distance between the transfer functions: $\left\|\tilde{\mathrm{H}}(\mathrm{s}) - \mathrm{H}(\mathrm{s})\right\|_\infty < \varepsilon$. To achieve this, BT transforms the original model into a ROM with a "balanced" state vector and then truncates the joint controllability-observability singular values of the system (so-called Hankel singular values) that sum up to the given threshold ε, as described in [2].

3 MOR by Balanced Truncation

3.1. Initial BT MOR

BT relies on the computation of the controllability Gramian P and observability Gramian Q, which are calculated as the solutions of the following Lyapunov matrix equations [2]:

$$(C^{-1}G)P + P(C^{-1}G)^T = -(C^{-1}B)(C^{-1}B)^T,\ (C^{-1}G)^T Q + Q(C^{-1}G) = -L^T L \qquad (4)$$

The main steps of the BT procedure are summarized in Algorithm 1. As can be seen, the operations involved (e.g., the solution of Lyapunov equations and the singular value decomposition [SVD]) are computationally expensive with complexity O(N3). Moreover, they are applied on dense matrices, since the Gramians P,Q are dense even if the system matrices C,G,B,L are sparse. Consequently, the significant computational and memory cost for deriving the ROM hinders the applicability of BT to large-scale models (with order N over a few thousand states).

Algorithm 1 MOR by balanced truncation

Input: $\mathbf{G}, \mathbf{C}, \mathbf{B}, \mathbf{L}$
Output: $\tilde{\mathbf{G}}, \tilde{\mathbf{C}}, \tilde{\mathbf{B}}, \tilde{\mathbf{L}}$

1: Solve the Lyapunov equations to obtain the Gramian matrices $\mathbf{P}$ and $\mathbf{Q}$ [6]
2: Compute the SVD of the Gramian matrices: $\mathbf{P} = \mathbf{U}_P \mathbf{\Sigma}_P \mathbf{V}_P^T$ and $\mathbf{Q} = \mathbf{U}_Q \mathbf{\Sigma}_Q \mathbf{V}_Q^T$
3: Find the square root of the Gramian matrices: $\mathbf{Z}_P = \mathbf{U}_P \mathbf{\Sigma}_P^{1/2}$ and $\mathbf{Z}_Q = \mathbf{U}_Q \mathbf{\Sigma}_Q^{1/2}$
4: Compute the SVD of the product of the roots: $\mathbf{Z}_Q^T \mathbf{Z}_P = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$
5: Compute transformation matrices: $\mathbf{T}_{(r\times N)} = \mathbf{\Sigma}_{(r\times r)}^{-1/2} \mathbf{U}_{(r\times N)} \mathbf{Z}_Q^T$, $\mathbf{T}^{-1}_{(N\times r)} = \mathbf{Z}_P \mathbf{V}_{(N\times r)} \mathbf{\Sigma}_{(r\times r)}^{-1/2}$
6: Compute ROM: $\tilde{\mathbf{G}} = \mathbf{T}_{(r\times N)} \mathbf{G} \mathbf{T}^{-1}_{(N\times r)}$, $\tilde{\mathbf{C}} = \mathbf{T}_{(r\times N)} \mathbf{C} \mathbf{T}^{-1}_{(N\times r)}$, $\tilde{\mathbf{B}} = \mathbf{T}_{(r\times N)} \mathbf{B}$, $\tilde{\mathbf{L}} = \mathbf{L} \mathbf{T}^{-1}_{(N\times r)}$

However, the products $(C^{-1}B)(C^{-1}B)^T$ *and* $L^T L$ have low numerical order compared to N, as p, *q << N*, resulting in low-rank Gramian matrices that can be approximated using low-rank techniques. This greatly reduces the complexity and memory requirements of the solution of the Lyapunov equations and the SVD analysis, which are now of order k instead of full order N.

3.2. Low-rank BT MOR

The essence of low-rank BT MOR is to iteratively project the Lyapunov Eq. (4) onto a lowerdimensional Krylov subspace (Kk) [7] and then solve the resulting small-scale equations to obtain low-rank approximate solutions of Eq. (4). In this work, we exploit the EKS to accelerate the convergence to the final solution [8]. The complete EKS method is presented in Algorithm 2.

Algorithm 2 Extended Krylov subspace method for low-rank solution of Lyapunov equations

Input: $\mathbf{G}_C \equiv \mathbf{C}^{-1}\mathbf{G}, \mathbf{B}_C \equiv \mathbf{C}^{-1}\mathbf{B}$ (or $\mathbf{G}_C^T, \mathbf{L}^T$)
Output: $\mathbf{Z}$ such that $\mathbf{P} \approx \mathbf{Z}\mathbf{Z}^T$

1: $j = 1$; $p = size\_col(\mathbf{B}_C)$
2: $\mathbf{K}^{(j)} = Orth([\mathbf{B}_C, \mathbf{G}_C^{-1}\mathbf{B}_C])$
3: **while** $j < maxiter$ **do**
4: $\quad \mathbf{A} = \mathbf{K}^{(j)T} \mathbf{G}_C \mathbf{K}^{(j)}$; $\quad \mathbf{R} = \mathbf{K}^{(j)T} \mathbf{B}_C$
5: $\quad$ Solve $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^T = -\mathbf{R}\mathbf{R}^T$ for $\mathbf{X} \in \mathbb{R}^{2pj \times 2pj}$
6: $\quad$ **if** converged **then**
7: $\quad\quad [\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}] = \mathbf{SVD}(\mathbf{X})$; $\quad \mathbf{Z} = \mathbf{K}^{(j)} \mathbf{U} \mathbf{\Sigma}^{1/2}$
8: $\quad\quad$ **break**
9: $\quad$ **end if**
10: $\quad k_1 = 2p(j-1)$; $k_2 = k_1 + p$; $k_3 = 2pj$
11: $\quad \mathbf{K}_1 = [\mathbf{G}_C \mathbf{K}^{(j)}(:, k_1 + 1 : k_2), \mathbf{G}_C^{-1} \mathbf{K}^{(j)}(:, k_2 + 1 : k_3)]$
12: $\quad \mathbf{K}_2 = Orth(\mathbf{K}_1)$ $\quad$ w.r.t. $\quad \mathbf{K}^{(j)}$
13: $\quad \mathbf{K}_3 = Orth(\mathbf{K}_2)$
14: $\quad \mathbf{K}^{(j+1)} = [\mathbf{K}^{(j)}, \mathbf{K}_3]$
15: $\quad j = j + 1$
16: **end while**

4 Experimental Evaluation

To evaluate the proposed MOR methods, we developed an EDA tool that implements the BT algorithms presented in Section 3. As depicted in Figure 1, the only input is a configuration file that defines the path to the MNA matrices along with some parameters. After applying BT MOR, the tool performs DC, transient, and SP analysis, to compare the ROM to the original model. The output includes the S-parameters and MNA matrices of the ROM. The cross-platform MORCIC tool was developed in C++ using the CMake automation sotware. All experiments were executed on a Linux workstation with a 3.60 GHz CPU and 16 GB of memory.

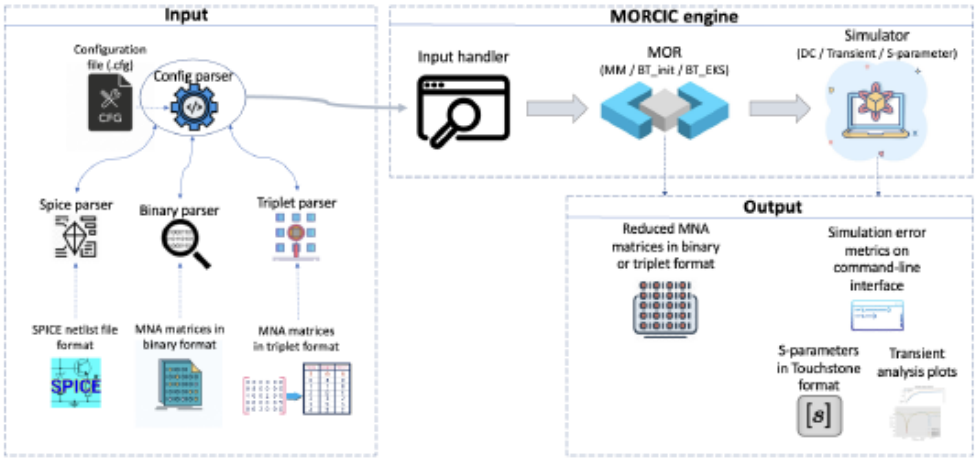

Figure 1: Software architecture of the MORCIC tool.

4.1 Initial BT MOR

For the evaluation of the initial BT MOR method, we used small-scale RC and RLCk models (i.e., real transmission lines) extracted by ANSYS RaptorX™ [3], which are presented in Table 1.

Table 1: Small-scale RC and RLCk models of transmission lines

Model	Initial order	#nodes	#resistors	#capacitors	#inductors	#mutual ind.	#ports
RC_1	48	48	202	273	0	0	2
RC_2	526	526	6667	6872	0	0	6
RLCk_1	5431	3084	2998	1282	2347	136271	2
RLCk_2	21800	12166	34635	31131	9634	23639237	6

As can be seen in Table 2, the mean relative error (MRE) and maximum relative error (MAX RE) for the DC analysis are lower than 0.61% and 0.62%, respectively. As for the SP analysis, MRE remains under 0.73% while MAX RE is below 1.89%. The reduction percentage to achieve accurate results is within 65-87%. However, considering the runtime and memory overhead of the initial method, its application on large-scale models is practically infeasible.

4.2 Low-rank BT MOR

To validate the accuracy and performance of the low-rank BT MOR method, we designed disparate circuits in the GlobalFoundries 22 nm FDSOI technology and extracted the corresponding large-scale RLCk models using RaptorX™ [3]. The choice of the benchmark circuits is driven by their diversity and therefore,

Table 2: Evaluation of ROMs generated by the initial BT MOR against the original models

Model	ROM order	Reduction (%)	DC analysis		SP analysis		Reduction time	Memory (GB)
			MRE (%)	MAX_RE (%)	MRE (%)	MAX_RE (%)		
RC_1	17	64.58	0.54	0.55	0.73	0.78	0.04 s	0.01
RC_2	93	82.32	0.09	0.17	0.32	0.43	2.51 s	0.09
RLCk_1	1131	79.18	0.61	0.62	0.25	1.89	1.12 h.	7.24
RLCk_2	2797	87.17	0.46	0.61	0.08	1.38	6 days	87.11

different metrics are used to describe their behavior. As shown in Table 3, the evaluated designs include Hybrid and Wilkinson couplers as well as typical transceiver blocks like low-noise-amplifiers (LNAs) and oscillators, where the metrics of interest are the reflection coefficients and performance (gain, noise, linearity). In our experiments, we utilized the MORCIC tool to generate ROMs with target accuracy comparable to RaptorX™.

Table 3: Large-scale RLCk models and metrics of interest for the designed circuits

Block/DUT	Initial Order	Ports	Mutual inductors	Simulated Metrics
Hybrid Coupler @28GHz	134710	5	79001243	S-parameters of coupler as:
Hybrid Coupler @56GHz	98024	5	52363149	power splitter & divider
Wilkinson Coupler @28GHz	129087	4	259462454	S-parameters of coupler as:
Wilkinson Coupler @56GHz	100888	4	193641938	power splitter & divider
VGA @28GHz	95189	13	40230583	S-parameters, attenuation
VCO @13GHz	104367	4	70445484	Spectrum, PN, osc. frequency
LNA Common-Source @56GHz	128574	9	72832315	S-parameters, gain, CP1dB,
LNA Cascode @28GHz	162881	11	98585323	IIP3, Noise Figure (NF)

The accuracy evaluation is performed by comparing the ROMs generated by the low-rank BT MOR against the reference ROMs obtained by RaptorX™, as the simulation of the original extracted models (i.e., full RLCk netlists) is infeasible. The evaluated metrics for the Hybrid and Wilkinson couplers at 28 GHz, both operating as power splitters, are demonstrated in Figure 2. As can be seen, the S-parameters of the MORCIC ROMs closely match those of the RaptorX™ ROMs across the frequency range, and most importantly at the frequency of interest. The insertion-loss error is lower than 0.5 dB, while the respective phases differ by less than 2 degrees.

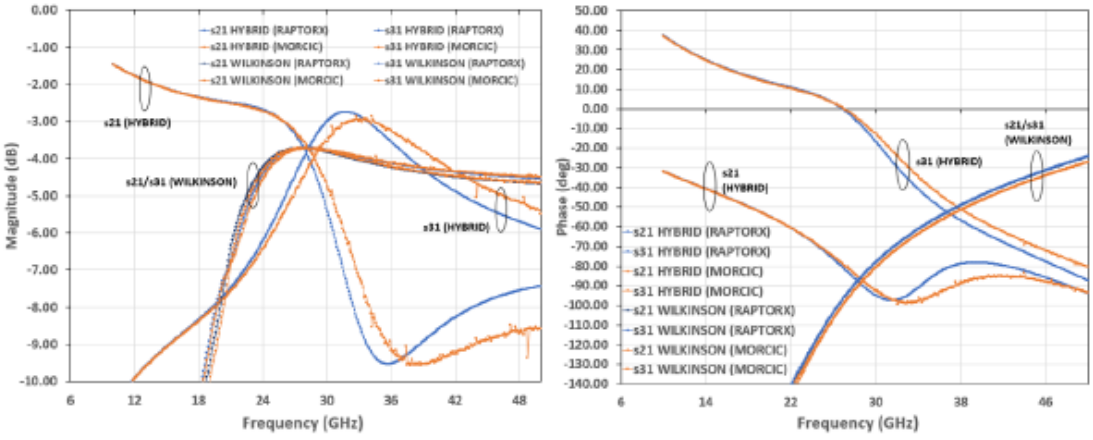

Figure 2: MORCIC vs RaptorX™ ROM accuracy: Hybrid/Wilkinson S-parameters and phases.

The efficiency of the MORCIC tool against RaptorX™ is demonstrated in Table 4. On average, MORCIC produces ×3.1 more compact ROMs. As the number of ports increases, the advantage of BT is more evident, with the maximum improvement in ROM order reaching ×5.5 for LNACasc 28. Although MORCIC has higher reduction time and memory requirements compared to RaptorX™, they are still reasonable and can be significantly improved in future work.

Table 4: MORCIC vs RaptorX™ ROM order and MOR performance

Model	Initial order	ROM order		Reduction time (s)		Memory (GB)	
		RaptorX™	MORCIC	RaptorX™	MORCIC	RaptorX™	MORCIC
VGA_28	95189	4744	1040	67	1037	5.6	29
Hybrid_56	98024	1267	397	104	613	3.7	44.1
Wilkinson_56	100888	765	320	154	570	3.8	45.1
VCO_13	104367	407	311	119	673	4.6	44.2
LNACS_56	128574	2172	716	74	1237	4.3	40.1
Wilkinson_28	129087	885	302	205	801	3.9	54.3
Hybrid_28	134710	787	399	217	1032	4	53.2
LNACasc_28	162881	4768	879	373	2866	15.7	73

5 Conclusions

In this paper, we present efficient BT MOR techniques to reduce electromagnetic RLCk models. The accuracy of the proposed methods has been evaluated across diverse benchmark circuits, such as the Hybrid/Wilkinson couplers, primarily comparing their S-parameters. Experimental results demonstrate that our low-rank BT MOR approach achieves sufficient accuracy while providing ROMs that are up to ×5.5 smaller than the ROMs obtained by ANSYS RaptorX™.

6 Acknowledgments

This research has been co-financed by the European Regional Development Fund and Greek national funds via the Operational Program ”Competitiveness, Entrepreneurship and Innovation,” under the call ”RESEARCH-CREATE-INNOVATE” (project code: T2EDK-00609).

7 References

[1] A. Odabasioglu et al., “Prima: Passive reduced-order interconnect macromodeling algorithm,” IEEE Trans. on CAD of Integrated Circuits and Systems, vol. 17, no. 8, pp. 645–654, 1998.

[2] S. Gugercin et al., “A survey of model reduction by balanced truncation and some new results,” International Journal of Control, vol. 77, no. 8, pp. 748–766, 2004.

[3] “Ansys-RaptorX.” [Online]. Available: www.ansys.com/products/semiconductors/ansys-raptorh

[4] C.-W. Ho et al., “The modified nodal approach to network analysis,” IEEE Trans. on Circuits and Systems, vol. 22, no. 6, pp. 504– 509, 1975.

[5] K. Grochenig, Foundations of Time-Frequency Analysis. Birkhauser, 2001.

[6] D. Lathauwer et al., “Computation of the canonical decomposition by means of a simultaneous generalized schur decomposition,” SIAM Journal on Matrix Analysis and Applications, vol. 26, no. 2, pp. 295–327, 2004.

[7] V. Simoncini, “A new iterative method for solving large-scale lyapunov matrix equations,” SIAM Journal on Scientific Computing, vol. 29, no. 3, pp. 1268–1288, 2007.

[8] C. Chatzigeorgiou et al., “Exploiting Extended Krylov Subspace for the Reduction of Regular and Singular Circuit Models,” in Proc. of the 26th Asia South Pacific Design Automation Conference, pp. 773–778, 2021.

[9] E. Bavier et al., “Amesos2 and Belos: Direct and Iterative Solvers for Large Sparse Linear Sys tems,” Sci. Program., vol. 20, no. 3, p. 241–255, jul 2012.

[10] D. Garyfallou et al., “A Combinatorial Multigrid Preconditioned Iterative Method for Large Scale Circuit Simulation on GPUs,” in Proc. of the 15th International Conference on Synthesis, Model ing, Analysis and Simulation Methods and Applications to Circuit Design, pp. 209–212, 2018.

[11] V. Simoncini, “A new iterative method for solving large-scale lyapunov matrix equations,” SIAM Journal on Scientific Computing, vol. 29, no. 3, pp. 1268–1288, 2007